ACEEE Int. J. on Signal and Image Processing , Vol. 5, No. 1, January 2014

Texture Unit based Monocular Real-world Scene Classification using SOM and KNN Classifier

N.P. Rath[1] and Bandana Mishra[2]
1Veer Surendra Sai University/Electronics and Communication Department, Sambalpur, India
Email: n_p_rath@hotmail.com
2Veer Surendra Sai University/Electronics and Communication Department, Sambalpur, India
Email: mishra18.bandana@gmail.com

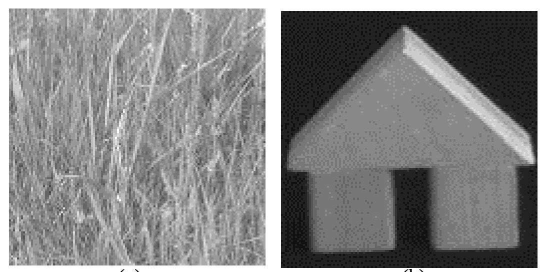
***Abstract* — In this paper a method is proposed to discriminate real world scenes in to natural and manmade scenes of similar depth. Global-roughness of a scene image varies as a function of image-depth. Increase in image depth leads to increase in roughness in manmade scenes; on the contrary natural scenes exhibit smooth behavior at higher image depth. This particular arrangement of pixels in scene structure can be well explained by local texture information in a pixel and its neighborhood. Our proposed method analyses local texture information of a scene image using texture unit matrix. For final classification we have used both supervised and unsupervised learning using K-Nearest Neighbor classifier (KNN) and Self Organizing Map (SOM) respectively. This technique is useful for online classification due to very less computational complexity.**

***Index Terms* - Image-depth, Texture unit, Texture unit matrix, scene image, Self Organizing Map (SOM), K-nearest neighbor classifier (KNN).**

I. Introduction

In monocular vision, image analysis is done by data derived from a single image in contrast to stereo vision, where spatial information is obtained by comparing different images of the same scene. Natural and manmade scene images exhibit a peculiar behavior with respect to variation in image depth, where depth of an image is the mean distance of the object from the viewer. Near manmade structures exhibit a homogenous and smooth view as shown in figure-1 (b). With increase in depth, smoothness of manmade image decreases because of inclusion of other artifacts. On the other hand 'near' natural scene is perceived as a textured region where roughness is high viz. figure-1(a). In 'far' natural scenes textured regions get replaced by low spatial frequency [9] components and give an appearance of smoothness. Such attributes of scene images can be perceived as follows: 'far' manmade and 'near' natural structures exhibit similar rough appearance and 'near' manmade and 'far' natural scenes exhibit smooth view and this textural difference can be explored to discriminate natural scenes and manmade scenes of similar depth.

Texture can be described as a repetitive pattern of local variations in image intensity. In a scene image, texture provides measures of some scene attributes like smoothness, coarseness and regularity. These extracted features constitute the feature vector which is analyzed further to classify scene images [1]. He and Wang [2] have proposed a statistical approach to texture analysis termed as 'Texture unit approach' which is used in our paper to distinguish between manmade and natural scenes. Here local texture information of a given pixel and neighborhood is characterized by 'Texture unit' and the unit is used further to quantify texture and to construct the feature vector. The feature vector is then subjected to KNN classifier and SOM classifier separately in order to obtain the final classification result. This is shown in the following block diagram (Fig 2).



(a) (b)

Figure 1. (a) Rough Image (b) Smooth Image

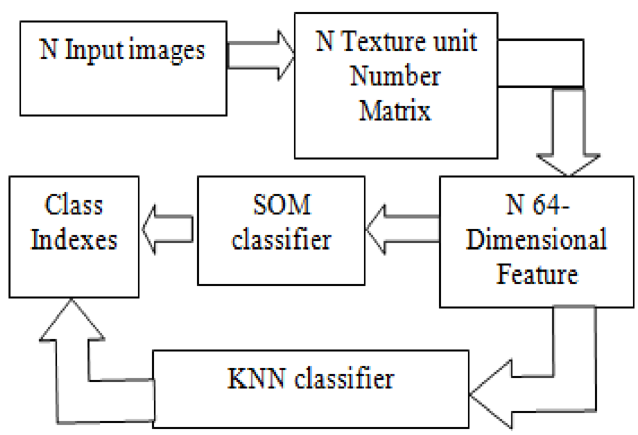


Figure 2. Block Diagram

Serrano et al [3] have proposed a method of scene classification where in the first level low-level feature sets such as color and wavelet texture features are used to predict multiple semantic scenes attributes and they are classified using support vector machine to obtain indoor/outdoor classification about 89%. In next level, the semantic scene attributes are then again integrated using a Bayesian network, and an improvised indoor/outdoor scene classification result of 90.7% is obtained. Raja et al [4] have proposed a method to classify the war scene category from the natural scene category. They have extracted Wavelet features from the images and feature vector is trained and tested using feed

DOI: 01.IJSIP.5.1.1402

forward back propagation algorithm using artificial neural networks and have reported classification success is 82%. Using the same database, Raja et al [5] have extracted features from images using Invariant Moments and Gray Level Co-occurrence Matrix (GLCM). They have reported that GLCM feature extraction method with Support Vector Machines classifier has shown result up to 92%. Chen et al [6] have proposed a scene classification technique where they have considered texture Unit Coding (TUC) concept to classify mammograms. The TUC generates a texture spectrum for a texture image and the discrepancy between two texture spectra is measured using information divergence (ID)-based discrimination criterion. They applied TUC along with ID classification mass in mammograms. Barcelo et al. [7] have proposed a texture characterization approach that uses the texture spectrum method and fuzzy techniques for defining 'texture unit boxes' which also takes care of vagueness introduced by noise and the different caption and digitations processes. Karkanis et al.[8] computed features based on the run length of the spectrum image representing textural descriptors of respective regions. They have characterized different textured regions within the same image, which is further applied successfully on endoscopic images for classifying between normal and cancer regions. Bhattacharya et al.[9] have used Texture spectrum concepts using 3x3 as well as 5x5 window for reduction in noise in satellite data. Al-Janobi[10] have proposed texture analysis method incorporating with the properties of both the gray-level co-occurrence matrix (GLCM) and texture spectrum (TS) methods. They have obtained Image texture information of an image using the method and they have worked on Brodatz's natural texture images. Chang et al. [11] have extended Texture Unit Coding (TUC) and proposed gradient texture unit coding (GTUC) where gradient changes in gray levels between the central pixel and its two neighboring pixels in a Texture unit (two pixels considered in the TUC), along with two different orientations is captured. Jiji et al.[12] proposed a method for segmentation of color texture image using fuzzy texture unit and color fuzzy texture spectrum. After locating color texture locally as well as globally segmentation operation is performed by SOMs algorithm. Rath et al. [13] have proposed a Gabor filter based scheme to segregate monocular scene images of real world natural scenes from manmade structures. Lee et al.[14] proposed a method for texture analysis using fuzzy uncertainty. They have introduced fuzzy uncertainty texture spectrum (FUTS), and it used as the texture feature for texture analysis. He Wang [15] have simplified the texture spectrum by reducing the 6,561 texture units into 15 units without significant loss of discriminating power. They have corroborated their claim by doing experimentation on Brodatz's natural texture images.

In this paper, a method is proposed by us where images of similar depth are classified to manmade and natural classes. In first stage of experiment scene images are converted to texture unit matrices and then feature vectors are generated from these matrices. In second stage, the feature vectors are subjected to SOM and KNN classifiers and classified results are obtained respectively. So in our method *'near'* and *'far'* scene images are getting classified to 'natural' and 'manmade' classes separately.

Brief outline of the paper is as follows. Section-II discusses the basic concepts of Base-3 Texture unit and its extended versions like Base-5, Base-7 and their threshold limits. This is followed by explanation of ordering way in texture unit. After that classifiers used in our work namely *Self Organizing Map* and *K Nearest Neighbor* are explained in brief. Section-III describes our experimental algorithm and section-IV presents elaborate discussion on experiments and results of the work. This paper is concluded in section-V discussing about possible implementation of our technique as a real time application.

II. Texture Unit Approach

He and Wang [1] have proposed a statistical approach to texture analysis termed as texture unit approach. Here local texture information for a given pixel and its neighborhood is characterized by the corresponding texture unit. It extracts the textural information of an image as it takes care of all the eight directions corresponding to its eight neighbors. In this work a neighborhood comprises of a 3x3 window taking the central pixel as image pixel.

A. Base3 texture unit number

In Texture unit approach a texture image can be decomposed into a set of essential small units called texture units. The neighborhood of 3x3 pixels which is denoted by a set V, comprising of nine elements:

$V = \{V_0, V1, V2, V3, V4, V5, V6, V7, V8\}$, where V_0: intensity value of the central pixel

V_1-V_8: intensity values of the neighboring pixels represented as V_i; i = 1, 2, 3…8

Then the corresponding texture unit can be represented as a set containing the elements,

TU = {E1, E2, . . ., E8}, where the elements of the texture unit Ei; i=1,2….8 are computed as follows:

$$E_i = \begin{cases} 0 \text{ if } v_i > v_o - \Delta \text{ and } v_i < v_o - \Delta \\ 1 \text{ if } v_i < v_o \\ 2 \text{ if } v_i > v_o \end{cases} \quad (1)$$

Δ = gray level tolerance limit.

Gray level tolerance limit is taken to obtain a distinguished response for textured and non textured region separately and this value is kept very small.

The intensity values V_i of 3x3 window are now replaced by the corresponding Ei. The TUN_{Base3} ranges from 0 to 6560. The texture unit number in base 3 is calculated as follows:

$$N_{TUBASE3} = \sum_{i=1}^{8} E_i 3^{i-1} = E_1 x 3^0 + E_2 x 3^1 + E_3 x 3^2 + E_4 x 3^3 + E_5 x 3^4 + E_6 x 3^5 + E_7 x 3^6 + E_8 x 3^7 \quad (2)$$

Where,

$N_{TUBase3}$: texture unit number with respect to Base-3.

DOI: 01.IJSIP.5.1.1402

ACEEE Int. J. on Signal and Image Processing , Vol. 5, No. 1, January 2014

E_i: i[th] element of texture unit set.
TU: {E1, E2,, . . .,E8}

B. Base-5 Texture Unit Matrix (TUM_{Base5})

The Base3 approach of texture units is unable to discriminate the differences from less or far-less and greater or far-greater with respect to the grey level value of central pixel. To incorporate this type of texture feature on a 3 x 3 window TUM_{Base5} and TUM_{Base7} approaches are proposed [16].

$$E_i = \begin{cases} 0 \text{ if } v_i > v_o - \Delta \text{ and } v_i < v_o + \Delta \\ 1 \text{ if } v_i < v_o \text{ and } v_i < X \\ 2 \text{ if } v_i < v_o \text{ and } v_i > X \\ 3 \text{ if } v_i > v_o \text{ and } v_i < Y \\ 4 \text{ if } v_i > v_o \text{ and } v_i > Y \end{cases} \tag{3}$$

Where X and Y are user-specified threshold limits, which is discussed elaborately in the following section.
Δ = gray level tolerance limit.

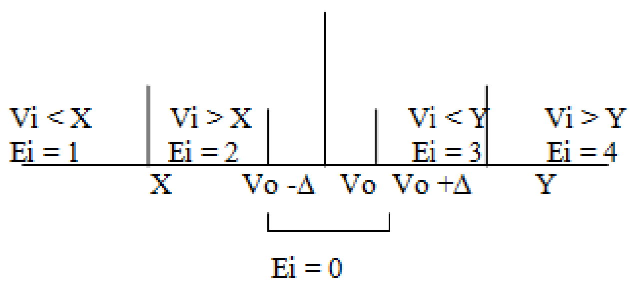


Figure 3. Base-5 Texture unit representation

Fig.3 is explaining the base-5 approach as per (3). The corresponding texture unit can be represented as a set containing eight elements, TU = {E1, E2, ... E8}. In this approach, (3) is used to determine the elements Ei of texture unit and Texture Unit number is computed using (4).In Fig. 4(a) 3x3 window is taken where central image pixel value is 140. Using (3) Texture unit is generated corresponding to each neighborhood pixel value; shown in Fig.4(b).Then an ordering way is chosen as discussed in sec[D] and Texture unit number is calculated. The TUN_{Base5} ranges from 0 to 2020. Minimum TUN_{Base5} value 0 is obtained by keeping all Ei values 0 in (4) and maximum TUN_{Base5} 2020 is obtained by keeping all Ei values 4 in (4).

$$N_{TUNBASE5} = \sum_{i=1}^{8} E_i 3^{(i-1)/2} = E_1 x\, 3^0 + E_2 x\, 3^{0.5} + E_3 x\, 3^1 + E_4 x\, 3^{1.5} + E_5 x\, 3^2 + E_6 x\, 3^{2.5} + E_7 x\, 3^3 + E_8 x\, 3^{3.5} \tag{4}$$

C. Base-7 Texture Unit Matrix (TUM_{Base7})

Similarly Base-7 approach of texture unit is proposed [16] where two threshold limits are taken. The range of TUN_{Base7} varies from 0 to 1172.

$$E_i = \begin{cases} 0 \text{ if } v_i > v_o - \Delta \text{ and } v_i < v_o + \Delta \\ 1 \text{ if } v_i < v_o \text{ and } v_i < X_l \text{ and } v_i < Y_l \\ 2 \text{ if } v_i < v_o \text{ and } v_i > X_l \text{ and } v_i < Y_l \\ 3 \text{ if } v_i < v_o \text{ and } v_i > X_l \text{ and } v_i > Y_l \\ 4 \text{ if } v_i > v_o \text{ and } v_i < X_u \text{ and } v_i < Y_u \\ 5 \text{ if } v_i > v_o \text{ and } v_i > X_u \text{ and } v_i < Y_u \\ 6 \text{ if } v_i > v_o \text{ and } v_i > X_u \text{ and } v_i > Y_u \end{cases} \tag{5}$$

Where X_l, Y_l, X_u, Y_u are user defined threshold limits.
Δ = gray level tolerance limit.
Texture unit number in base-7 is computed as:

$$TUN_{BASE5} = \sum_{i=1}^{8} E_i X 7^{(i-1)/3} \tag{6}$$

90	130	145
160	140	200
100	140	250

(a)

1	2	0
3		4
1	0	4

(b)

[1 2 0 4 4 0 1 3] → 1114

(c)

Figure 4. Base-5 technique (a) 3x3 window (b) Texture unit (c) Texture unit to Texture unit Number

D. Decision regarding Threshold Limits

The user defined threshold values in Base-5 are X and Y. In this paper following strategy is adopted to determine their values. For each 3x3 window, maximum pixel value (max), minimum pixel value (min) among all the nine elements and central pixel value (Vo) are computed. X and Y are the mid points of (Vo, min) and (Vo, max) respectively as given in (7).

$$X = V_o - \frac{V_o - \min}{2}, Y = V_o + \frac{\max - V_o}{2} \tag{7}$$

Similarly for Base-7 approach the threshold values are X_l, Y_l, X_u and Y_u are chosen between (Vo , min) and (Vo, max) as shown in (8).

$$X_l = \min + \frac{(V_o - \min)}{3}, Y_l = V_o - \frac{(V_o - \min)}{3}$$
$$X_u = V_o + \frac{(\max - V_o)}{3}, Y_u = \max - \frac{(\max - V_o)}{3} \tag{8}$$

E. Ordering way of Texture Unit number

Ordering way of a texture unit is to arrange the texture unit for a 3x3 window; fig 4(b) (comprising of 8 elements for a 3x3 window fig 4(a)). This box may be arranged in maximum 8

DOI: 01.IJSIP.5.1.1402

possible ways starting from each individual element giving rise to 8 possible ordering ways for a window. Thus any window will provide 8 texture unit numbers for an image pixel. In fig. 5, an example of a texture unit is given along with 8 possible ordering ways and their corresponding TUN. Fig. 6 displays a synthetic texture image and its Texture unit matrices in base-5 for 8 ordering ways.

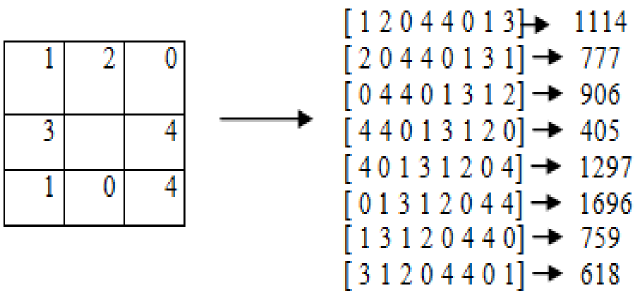


Figure 5. Example showing 8 ordering ways and 8 Texture unit numbers

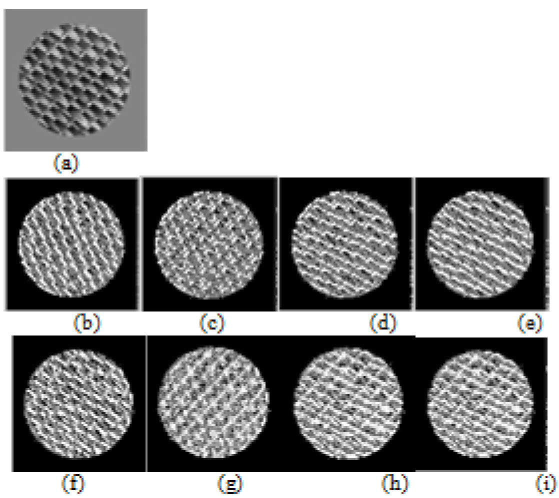


Figure 6. Texture image (b) – (h) Texture unit matrix for base-5 ordering way 1 to 8

F. Self Organizing Map

Self-organizing feature maps (SOFM) learn to classify input vectors according to how they are grouped in the input space in an unsupervised way [17]. when input is presented to the first layer, it computes the distance between input vector and weights associated with the neurons. in an iteration the distances from each input are compared using compete transfer function at the output layer and winner neuron is decided. Winning neuron gets value 'one' and others get 'zero'. Weight of winner neuron is updated using kohonen rule and subjected to next iteration. In this way specified number of iterations is performed and final classification result is obtained.

G. K Nearest Neighbor classifier

Nearest-neighbor methods [18] work better with data where features are statistically independent. This is because nearest-neighbor methods are based on some form of distance measure and nearest-neighbor detection of test data is not dependent on their feature interaction. Out of all training samples, k nearest neighbors to the test sample are chosen, where, value of k is user-defined. The average distance of classes from test data; based on the distance from the test data of all training samples of given classes found within the domain. The class with the smallest distance from the test data is declared the winner and the test pattern is allocated to this class.

- Out of n training vectors, identify k nearest-neighbors, *irrespective* of class label. k is chosen to be odd.
- Out of these k samples, identify the number of vectors, k_i, those belong to class w_i, i=1, 2,3,…M (M stands for total number of classes). Obviously $\sum_i k_i = k$.
- Find the average distance d_i that represents the distance between test pattern $X = \{X_1, X_2, \ldots, X_N\}$ and k_i nearest-neighbors (Y_i, which is a N-dimensional vector) found for class w_i, i=1,2,…, M. then distance between test pattern X and class W_i is
- $$d_i = \frac{1}{k_i}\sum_{i=1}^{k_i}\sum_{j=1}^{N}\left|(X_{ij} - Y_{ij})\right|$$
- Assign X to class C if its d_i is smallest, i.e. $X \in W_c$ if $d_c < d_i,\ \forall i$, such that $C \in [1,...M]$ ..

The decision in this model does not depend on the number of nearest-neighbors found but solely on the average distance between the test pattern and samples of each class found.

III. Algorithm to Compute Feature Vector

The feature vectors for Base-3, Base-5 and base-7 approaches are computed as follows. Appropriate equations are used for respective TUN techniques:

- 150 Scene images of size 128x128 are considered.
- 3 x 3 Window is chosen taking each image-pixel as its central pixel.
- To process the border pixels, image is padded with 2 rows and 2 columns making its size 130x130.
- 8 texture units (TU) for all possible ordering ways are generated with respect to each window.
- Each Texture Unit (TU) is then converted to Texture unit number (TUN). Thus each image pixel corresponds to 8 TUNs.
- Above process is repeated for entire image pixels. Each image pixel is replaced by its corresponding TUN, which results in 8 Texture unit matrices of size 128 x 128.
- Feature matrix (128x128) is obtained by taking pixel wise minimum of above 8 TUN matrices.
- Then feature matrix is down sampled to 8 x 8 matrix and resized into 64 x 1 column vector representing the feature vector of an image.

DOI: 01.IJSIP.5.1.1402

IV. Experimental Results And Analysis

Scene images of manmade and natural scenes are obtained from [19], [20]. Sample images belonging to all the four classes are shown in fig. 12 and fig. 13. It has been observed that 'manmade *'near'* images are smoother than 'manmade *'far'* images. Similarly natural *'far'* images appear smoother than natural *'near'* images. So we have segregated our database of 300 images into two databases (each 150 images) such as *'near'* image database and *'far'* image database which is accomplished according to human perception. For *'near'* image database we have considered natural scene images of bush leaves etc. and manmade scenes like toys, house interior as *'near'* scene images having depth within 10 meters. Similarly for *'far'* image database we have considered natural scene images of open field, panoramic views, mountains and manmade scenes of inside city views, tall building and urban views having depth about 500 meters. In our work SOM and KNN classifiers are used to classify these scene images. They have classified the scene images of similar depth into natural and manmade categories and their performance has been explained.

Images taken in this work are of size 128 x 128. In an image, each image pixel is taken as central pixel and a 3X3 window is selected around that pixel. We found that a 3X3 window captures texture variations better than large window size. In each window all angle orientations (0^0, 45^0, 90^0, 135^0, 180^0, 225^0, 270^0, 325^0) are considered as eight neighbors. Then texture unit array is generated, where we have chosen grey level tolerance limit $\Delta = 5$. Eight texture units (TU) are obtained for a window corresponding to 8 possible ordering ways. Then each TU is converted to its corresponding TUN, which replaces the central image pixel. Thus one image pixel gives rise to eight TUNs and subsequently the entire image gives rise to 8 TUN matrices of size 128X128 which is same as image size. We have experimented and inferred that pixel wise minimum of the Texture unit matrices corresponding to all ordering ways gives best results. Therefore pixel wise minimum of these eight matrices is obtained which constitutes the feature matrix of size 128x128.Feature matrix is down sampled to 8x8 matrix and resized to 64x1 column vectors which is the feature vector of an image. The process is repeated to produce 150 (64-dimensonal) feature vectors for each database. Output responses (Base-5 Texture unit matrix) of some scene images are displayed in fig-7

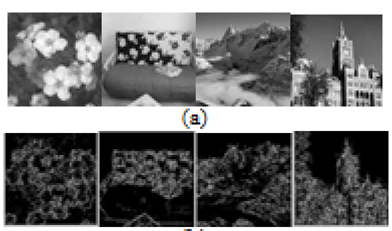
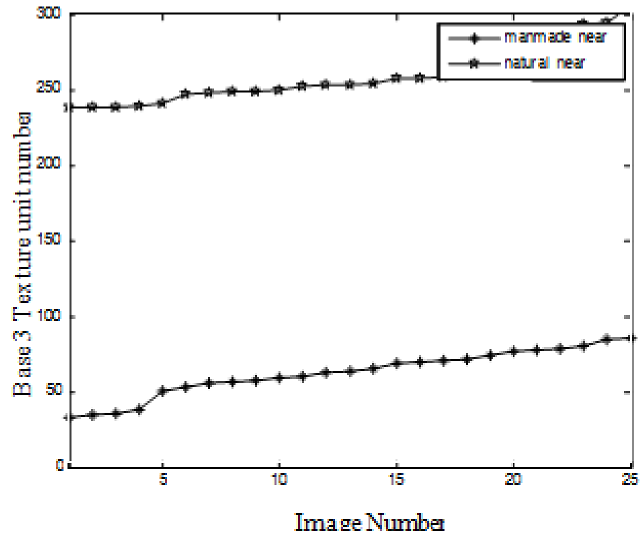
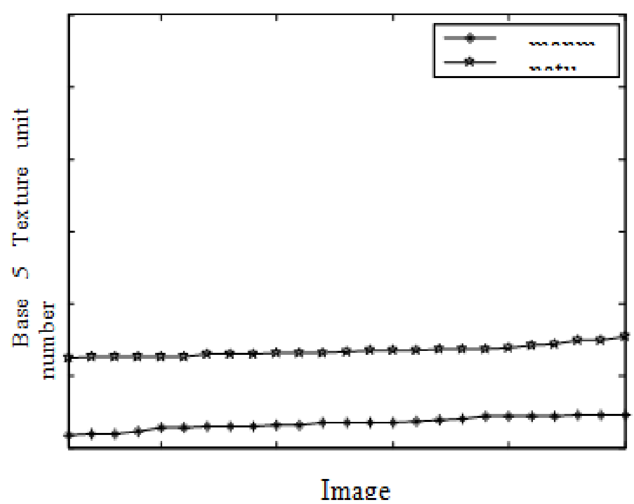


Figure 7. Texture unit matrix outputs (b) of some specimen images (a)

We have experimented on Base-3, Base-5 and Base-7 methods of TUN. As shown in fig.-8, fig.-9, graph is plotted between TUN and number of image. TUN values for 50 *'near'* images (25 from each category) are plotted in figure 8(a, b, c) for base-3, base-5 and base-7 respectively. Manmade *'near'* images are shown with star marker and natural *'near'* images are shown with pentagon marker. From the graphs it is observed that in *'near'* image database inter class gap between manmade *'near'* and Natural *'near'* is very distinct in base-3 (Fig. 8.a) technique than that of base-5 (Fig. 8.b) and base-7 (Fig. 8.c). In *'far'* image database similar behavior is noticed. We have concluded from the graphs that all the above techniques are showing distinguishable inter class gap between natural and manmade scenes. But Base-3(Fig. 9.a) shows better inter class gap in comparison to base-5 (Fig. 9.b) and base-7 (Figure 9.c). Therefore we have employed base-3 technique while using KNN classifier.



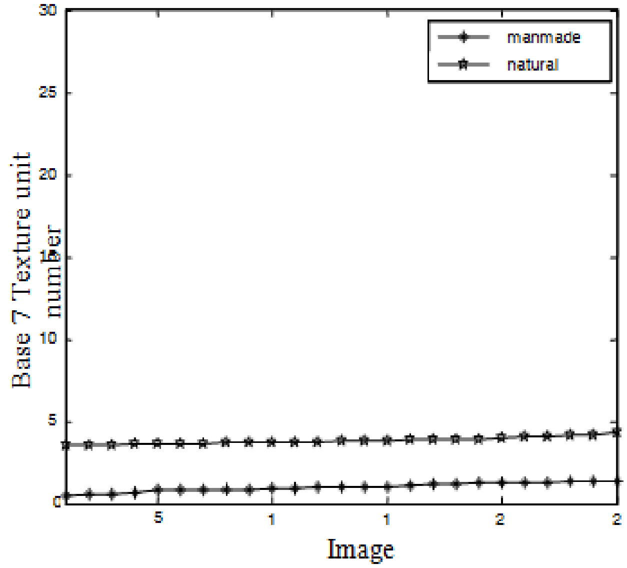
(a)



(b)

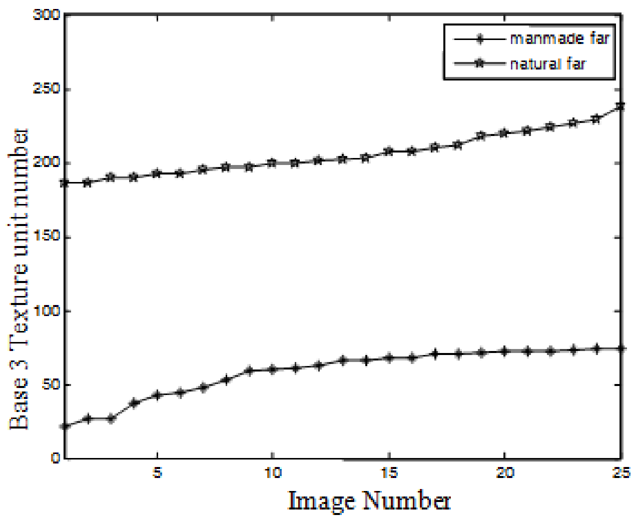
We have classified the scene images using K-Nearest Neighbor classifier (*viz.* section-II (*G*)). To classify both *'far'* and *'near'* scenes, initially 100 (64 dimensional) feature vectors are given as training vectors to KNN. The number of output classes (i) is two (manmade and natural). The network is tested for various values of k with 50 test images. Table -2

DOI: 01.IJSIP.5.1.1402

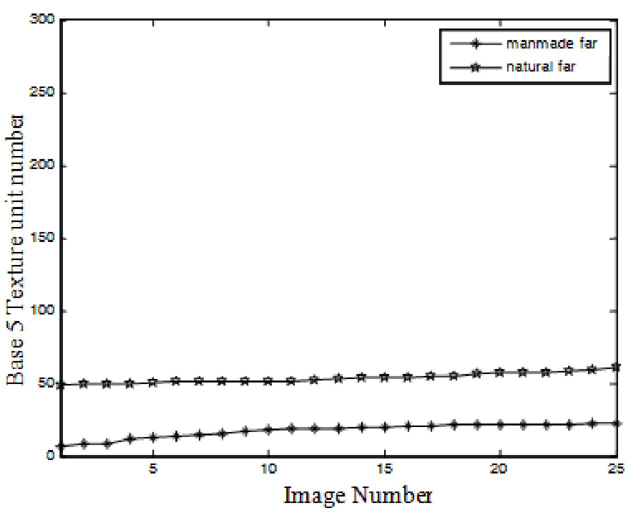


(c)

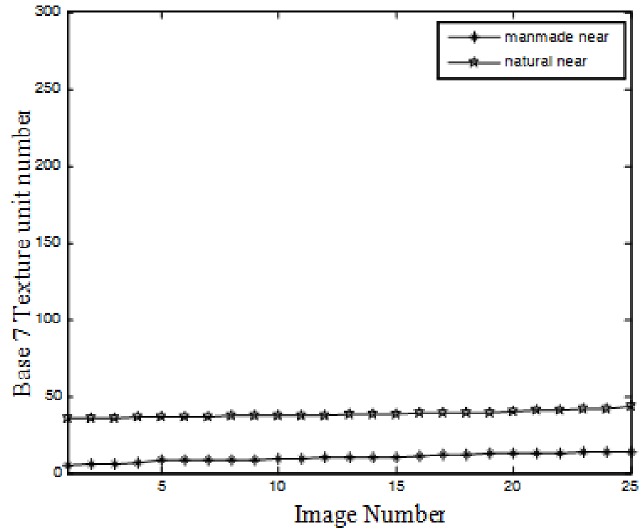
Figure 8. *'near'* image database performance in base-3(a); base-5(b); base-7(c)



(a)



(b)



(c)

Figure 9. *'far'* image database performance in base-3(a); base-5(b); base-7(c)

shows the number of misclassifications found for various k values. It is observed that for k = 1, number of misclassifications starts at a lower value and with increasing k it increases. But at higher values of k, number of misclassification go on decreasing and it attains a nearly constant value i.e. further increase in k does not alter the number of misclassifications. This behavior of k is verified for all the TUN approaches i.e. Base-3, Base-5, Base7. Figure 10 depicts the behavior of number of misclassifications with varying values of k using base-3 method.

Results obtained in KNN classification is as follows. Using Base-3 approach, when value of k = 5, number of misclassifications is found to be 1 for *'near'* database and 2 for *'far'* database for 50 test images. Thus we obtained classification result of 98% for *'near'* database and 96% for *'far'* database. At this value of k it is observed that both the databases respond to KNN classifier with minimum number of misclassification in base-3 approach as compared to other approaches like base-5 and base-7. It is also observed that misclassifications occur where structure of the scene is found to be ambiguous.

Again same set of images are given to Self organizing Map classifier (*viz.* section-II (*F*)) to classify the scene images. To classify the *'near'* scenes, 150 (64 dimensional) feature vectors are given as inputs to SOM. The number of output classes is two (manmade and natural). The network is trained for 500 iterations as we have found that higher number of iterations does not improve the result. Results obtained using SOM classifier is 98% for *'near'* database where number of misclassified images was found to be 3. Similarly for *'far'* scenes we obtained a classification result of 96% where number of misclassified images was found to be 6.

In table-I we have compared the misclassified results for base-3, base-5 and base-7 approaches for *'near'* and *'far'* databases. It is found that number of misclassifications found in case of base-5 is minimum in comparison to base-3 and base-7.

DOI: 01.IJSIP.5.1.1402

TABLE I. COMPARISON OF BASE-3, BASE-5, BASE-7 APPROACH USING 150 IMAGES IN EACH DATABASE

Approach	No. of misclassifications	
	'near' Images	*'far' Images*
Base-3	6	8
Base-5	3	6
Base-7	3	7

TABLE II. NO OF MISCLASSIFICATION FOR VARIOUS VALUES OF K IN NEAR AND FAR IMAGE DATABASE IN BASE-3, BASE-5, BASE-7 METHODS

k	Base3		Base5		Base7	
	Near	*Far*	*Near*	*Far*	*Near*	*Far*
1	1	2	1	2	2	3
5	1	2	2	2	2	4
11	2	6	1	8	1	9
15	4	9	2	8	1	7
21	5	4	3	9	2	12
25	3	4	3	6	2	9
31	5	2	4	5	3	8
35	6	2	4	6	4	8
41	7	1	2	4	3	5
45	6	3	3	4	2	6
51	5	3	3	4	1	6
55	3	4	1	4	2	4
61	3	4	1	4	2	5
65	3	4	2	4	1	4
71	1	3	1	3	1	3
75	1	3	1	3	1	3
81	1	3	1	3	1	4
85	1	3	1	3	1	3

Misclassification occurs for the images having ambiguous structure i.e. images whose structure are not well defined. Some of the misclassified images are shown in fig.-11 the near natural scene shown in fig.11 (a) is having smooth texture due to the flower petals, thus misclassified as near manmade scene. Fig.-11 (b) is misclassified as a near natural scene because roughness is high in the global image-structure, whereas fig.-11 (c) shows a misclassified image because of low illumination.

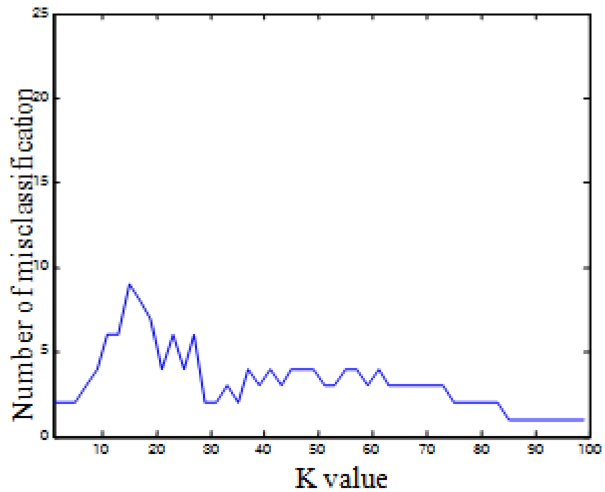


Figure 10. Variations of k with number of misclassification

(a) (b) (c)

Figure 11. Misclassified images

Figure 12. Sample Images belonging to 2 classes; *'near'* natural scene (Row1);*'near'* manmade scene (Row-2)

Figure 13. Sample Images belonging to 2 classes; 'far' Natural scene (Row-1);'far' manmade scene (Row-2)

CONCLUSIONS

In this work, we have proposed a method where the textural information of a monocular scene image is captured

by Texture unit matrix to analyze the scene images. Monocular scene images require simpler image acquisition systems and computationally more efficient in terms of execution time. Our algorithm is capable to capture the depth information similar to human perception from single view scene images. It shows that spatial properties are strongly co-related with the spatial arrangement of structures in a scene. It may be used to locate war scenes (for example scenes having war tanks, ammunition etc.) from natural scenes. We found that obtaining the Texture unit matrix with respect to base-5 and taking pixel wise minimum of all ordering ways, produced the best result *i.e*. 98% for near scene image database and 96% for far scene image database. This method may be explored further for automated classification of scene images irrespective of image-depth. The technique developed here for assessing depth information may be employed for localization of objects with depth as a cue. It may be useful for online classification due to very less computational complexity.

References

[1] Srinivasan, G. N., Shobha, G.: Statistical Texture Analysis. in proceedings of world academy of science, engineering and technology. Vol. 36 ISSN pp.2070-3740(2008).

[2] He, D.C., Wang,L.: Texture unit, texture spectrum and texture analysis. IEEE Trans. Geoscience and Remote Sensing. vol. 28 (4) pp.509–512(1990).

[3] Serrano, Navid., Savakis, E., Andreas., Luo, Jiebo.,: Improved scene classification using efficient low-level features and semantic cues. Pattern Recognition Society Vol. 3 (3) pp. 1773-1784 (2004)

[4] Raja, Daniel Madan, S., Shanmugam, A.: Wavelet Features Based War Scene Classification using Artificial Neural Networks. International Journal on Computer Science and Engineering Vol. 2 (9) pp. 3033-3037(2010)

[5] Raja, Daniel Madan, S., Shanmugam, A.: ANN and SVM Based War Scene Classification Using Invariant Moments and GLCM Features: A Comparative Study. International Journal of Machine Learning and Computing vol. 2(6) pp. 869-873(2012)

[6] Chen, Yuan., Chang, Chein-I.,: A New Application of Texture Unit Coding to Mass Classification for Mammograms; International conference on image processing vol.9, pp. 3335-3338 (2004)

[7] Barcelo.A,Montseny,E.,Sobrevilla,P.: Fuzzy Texture Unit and Fuzzy Texture Spectrum for texture characterization. Fuzzy Sets and Systems vol.158 pp.239 – 252(2007).

[8] Karkanis.S.,Galoousi.K.,Maroulis.D.: Classification of endoscopic images based on Texture Spectrum. In Proceedings of Workshop on Machine Learning in Medical Applications, Advance Course in Artificial Intelligence-ACAI(1999)

[9] Bhattacharya,K.,Srivastava,P.K.,Bhagat,A.; A Modified Texture filtering technique For Sattelite images. 22nd Asian conference on remote sensing, (2001).

[10] Al-Janobi, Abdulrahman.: Performance evaluation of cross-diagonal texture. matrix method of texture analysis. Pattern Recognition vol.34 pp.171-180(2001)

[11] Chang,Chein-I,Chen Yuan,: Gradient texture unit coding for texture analysis. Optical Engineering Vol.43(8) pp.1891–1903(2004)

[12] Jiji,G.W., Ganesan,L.: A new approach for unsupervised segmentation. Applied Soft Computing Vol.10 pp.689–693(2010)

[13] Rath, N.P., Mandal. M., Mandal, A. K.: Discrimination of manmade structures from natural scenes using Gabor filter. International conference on intelligent signal processing and robotics, Indian Institute of Information Technology, Allahabad, India, 20-23,(2004)

[14] Lee,Y.G.,Lee,G.H.,Hsueh,Y.C.: Texture classification using fuzzy uncertainty texture spectrum. Neuro-computing Vol.20 pp.115-122(1998)

[15] He,D.C.,Wang,L.:Simplified Texture Spectrum for Texture Analysis. Journal of Communication and Computer, ISSN 1548-7709, USA(2010).

[16] Wiselin Jiji G. and Ganesan L.,: A New Approach for Unsupervised segmentation and Classification using Color Fuzzy Texture Spectrum and Doppler for the Application of Color Medical Images: Journal of Computer Science,02(1),pp. 83-91,2006

[17] Kohonen,T. Self-Organization and Associative Memory, 2nd Edition, Berlin:Springer-Verlag,1987

[18] S. Singh, J. Haddon, M Markou, Nearest-neighbor classifier in Natural Scene analysis, Pattern Recognition, vol.34, pp. 1601-1612(2001).

[19] A. Oliva, A. Torralba, Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope,Int'l J. Computer Vision, vol. 42(3) pp. 145-175(2001).

[20] L. Fei-Fei ,P. Perona, A Bayesian Hierarchical Model for Learning Natural Scene Categories,Proc. IEEE CS Conf. Computer Vision and Pattern Recognition, pp. 524-531 (2005).

DOI: 01.IJSIP.5.1.1402